

(19)



JAPANESE PATENT OFFICE

PATENT ABSTRACTS OF JAPAN

(11) Publication number: 10162011 A

(43) Date of publication of application: 19 . 06 . 98

(51) Int. Cl

G06F 17/30
G06F 13/00

(21) Application number: 08316305

(71) Applicant: TOSHIBA CORP

(22) Date of filing: 27 . 11 . 96

(72) Inventor: KURODA KAZUYO

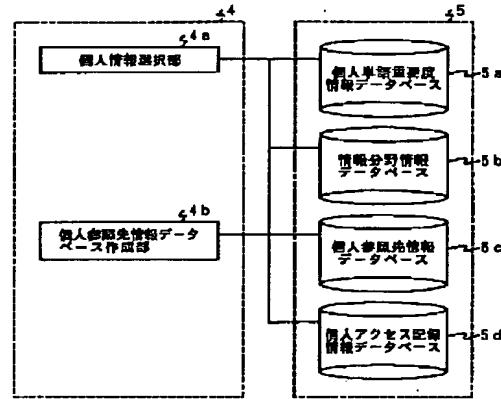
(54) **INFORMATION RETRIEVAL METHOD,
INFORMATION RETRIEVAL SYSTEM,
INFORMATION RETRIEVAL TERMINAL
EQUIPMENT, AND INFORMATION RETRIEVAL
DEVICE**

(57) Abstract:

PROBLEM TO BE SOLVED: To supply necessary network information in a computer system connected to a network to user.

SOLUTION: An individual reference destination information data base generation part 4b stores access recording to the computer system in an individual access recording information data base 5d and stores the address of network information which is frequently accessed by referring to the recording in an individual reference destination information data base 5c. An individual information selection part 4a analyzes network information stored in the data base 5c, generates an individual word significance information data base 5a, refers to the information field information data base 5b, and selects and supplies network information fitted to a user.

COPYRIGHT: (C)1998,JPO



(19)日本国特許庁 (JP)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平10-162011

(43)公開日 平成10年(1998)6月19日

(51)Int.Cl.⁶
G 0 6 F 17/30
13/00

識別記号
3 5 5

F I
G 0 6 F 15/403
13/00
15/40
15/403
3 4 0 A
3 5 5
3 1 0 F
3 4 0 B

審査請求 未請求 請求項の数11 O.L (全 12 頁)

(21)出願番号 特願平8-316305

(22)出願日 平成8年(1996)11月27日

(71)出願人 000003078
株式会社東芝
神奈川県川崎市幸区堀川町72番地

(72)発明者 黒田 和代
東京都青梅市末広町二丁目九番地 株式会
社東芝青梅工場内

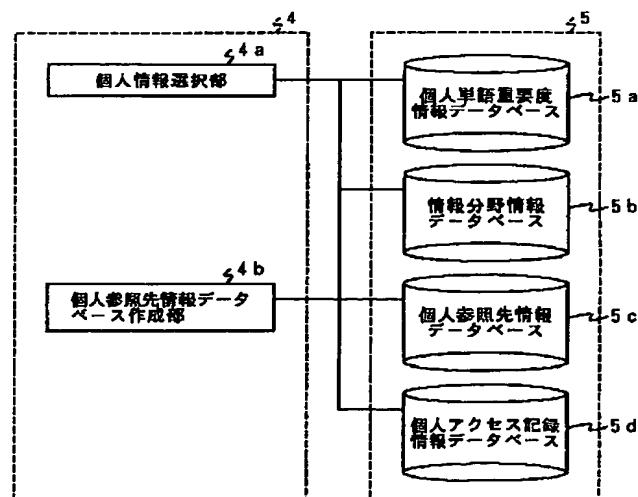
(74)代理人 弁理士 三澤 正義

(54)【発明の名称】 情報検索方法、情報検索システム、情報検索端末装置および情報検索装置

(57)【要約】

【課題】 本発明はネットワークに接続されたコンピュータシステム内の必要なネットワーク情報をユーザーに提供することに関するものである。

【解決手段】 個人参照先情報データベース作成部4 bによりコンピュータシステムへのアクセス記録を個人アクセス記録情報データベース5 dに格納し、この記録を参照して頻繁にアクセスするネットワーク情報のアドレス等を個人参照先情報データベース5 cに格納する。個人情報選択部4 aは、データベース5 cに格納されたネットワーク情報を解析して個人単語重要度情報データベース5 aや情報分野情報データベース5 bを参照してユーザーに適したネットワーク情報を選択、提供する。



【特許請求の範囲】

【請求項1】 ネットワークと、
 このネットワークに接続している複数のコンピュータシステムと、
 各コンピュータシステム内のネットワーク情報を格納するネットワーク情報格納手段と、
 アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納する手段と、
 アクセス記録情報を基にして頻繁にアクセスしたネットワーク情報の題目を格納する参照先情報格納手段と、
 ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納する手段と、
 参照先情報格納手段に格納されているネットワーク情報の題目から重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、決定した情報分野を基にしてネットワーク情報格納手段に格納されているネットワーク情報の中から所定のネットワーク情報を選択する手段と、
 選択したネットワーク情報を提供する手段とを有することを特徴とする情報検索システム。

【請求項2】 ネットワークと、
 このネットワークに接続している複数のコンピュータシステムと、
 各コンピュータシステム内のネットワーク情報を格納するネットワーク情報格納手段と、
 アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納する手段と、
 アクセス記録情報を基にして頻繁にアクセスしたネットワーク情報のアドレスとその題目を格納する参照先情報格納手段と、
 ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納する手段と、
 参照先情報格納手段に格納されているネットワーク情報の題目のみから単語を抽出するか、ネットワーク情報の題目およびそのアドレスに所在する情報本体から単語を抽出するか、またはネットワーク情報の題目、そのアドレスに所在する情報本体およびこの情報本体から参照された他のネットワーク情報の情報本体から単語を抽出し、抽出した単語にその出現場所や出現回数を基にして重要度点を付けることによって重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、決定した情報分野を基にしてネットワーク情報格納手段に格納されているネットワーク情報の中から所定のネットワーク情報を選択する手段と、
 選択したネットワーク情報を提供する手段とを有することを特徴とする情報検索システム。

【請求項3】 ネットワークと、

このネットワークに接続している複数のコンピュータシステムと、
 各コンピュータシステム内のネットワーク情報を格納するネットワーク情報格納手段と、
 アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納する手段と、
 アクセス記録情報を基にして頻繁にアクセスしたネットワーク情報のアドレスとその題目を格納する参照先情報格納手段と、
 ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納する手段と、
 参照先情報格納手段に格納されているネットワーク情報の題目から重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、決定した情報分野を基にしてネットワーク情報格納手段に格納されているネットワーク情報の中から所定のネットワーク情報を選択する手段と、
 選択したネットワーク情報を提供する手段と、
 ユーザの指定によりネットワーク情報のアドレスとその題目とを参照先情報格納手段に格納させ、前記アクセス記録情報を参照し、アクセス回数が所定期間内に所定回数以上であるネットワーク情報のアドレスとその題目を参照先情報格納手段に格納させ、アクセス回数が所定期間内に所定回数未満であるネットワーク情報のアドレスとその題目とを参照先情報格納手段から削除させる手段とを有することを特徴とする情報検索システム。

【請求項4】 ネットワークに接続している複数のコンピュータシステム内のネットワーク情報を格納し、
 アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納し、
 アクセス記録情報を基にして頻繁にアクセスしたネットワーク情報のアドレスとその題目を格納し、
 ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納し、
 格納されているネットワーク情報の題目から重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、
 決定した情報分野を基にして格納しているネットワーク情報の中から所定のネットワーク情報を選択し、
 選択したネットワーク情報を提供することを特徴とする情報検索方法。

【請求項5】 ネットワークに接続している複数のコンピュータシステム内のネットワーク情報を格納し、
 アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納し、
 アクセス記録情報を基にして頻繁にアクセスしたネット

50

ワーク情報のアドレスとその題目とを格納し、ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納し、格納されているネットワーク情報の題目のみから単語を抽出するか、ネットワーク情報の題目およびそのアドレスに所在する情報本体から単語を抽出するか、またはネットワーク情報の題目、そのアドレスに所在する情報本体およびこの情報本体から参照された他のネットワーク情報の情報本体から単語を抽出し、抽出した単語にその出現場所や出現回数を基にして重要度点を付けることによって重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、決定した情報分野を基にして格納しているネットワーク情報の中から所定のネットワーク情報を選択し、選択したネットワーク情報を提供することを特徴とする情報検索方法。

【請求項6】アドレスを指定してネットワークに接続している複数のコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納する手段と、格納したアクセス記録情報を基にして頻繁にアクセスしたネットワーク情報の題目を格納する手段と、ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納する手段と、格納したネットワーク情報の題目から重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、決定した情報分野を基にして格納されているネットワーク情報の中から所定のネットワーク情報を選択し、選択したネットワーク情報を取得する手段を有することを特徴とする情報検索端末装置。

【請求項7】ネットワークに接続している複数のコンピュータシステム内のネットワーク情報を格納するネットワーク情報格納手段と、アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報の中で頻繁にアクセスされたネットワーク情報の題目から重要度の高い単語が抽出され、予め単語に対応付けられた情報分野を参照して重要度の高い単語が属する情報分野が決定されることにより、この情報分野を基にしてネットワーク情報格納手段に格納しているネットワーク情報の中から選択された所定のネットワーク情報を提供する手段とを有することを特徴とする情報検索装置。

【請求項8】前記ネットワーク情報は、テキスト情報、画像情報および音声情報のうち少なくとも1つを含むことを特徴とする請求項1乃至3のいづれか1項に記載の情報検索システム。

【請求項9】前記ネットワーク情報は、テキスト情報、画像情報および音声情報のうち少なくとも1つを含

むことを特徴とする請求項4又は5に記載の情報検索方法。

【請求項10】前記ネットワーク情報は、テキスト情報、画像情報および音声情報のうち少なくとも1つを含むことを特徴とする請求項6に記載の情報検索端末装置。

【請求項11】前記ネットワーク情報は、テキスト情報、画像情報および音声情報の少なくとも1つを含むことを特徴とする請求項7に記載の情報検索装置。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】本発明は、インターネットのような大規模ネットワークに接続している多数のコンピュータシステム（サーバー）内の情報（ネットワーク情報）をユーザ各個人に適した内容でかつ適した分量でユーザーに提供するための情報検索方法、情報検索システム、情報検索端末装置及び情報検索装置に関する。

【0002】

【従来の技術】インターネットのような大規模ネットワークに接続している多数のコンピュータシステム（サーバー）内の情報（ネットワーク情報）の中からユーザ各個人に必要なネットワーク情報を提供するために、ユーザーが能動的にネットワーク情報を検索することなくユーザーに適した内容でかつ適した分量の情報をユーザーに提供するという情報フィルタリングする情報検索システムが近年開発されている。ネットワーク情報には、テキスト情報、画像情報、音声情報等の少なくとも1つの情報が含まれる。インターネットでは、これらの情報を用いてユーザーのホームページが作成される。

【0003】ある情報検索システムでは、各コンピュータシステム内のネットワーク情報を収集し、収集したネットワーク情報を情報検索システムによって管理されている検索用データベースに格納する。ユーザーによって情報検索のための検索キーワードが入力された場合、入力されたキーワードを基にしてキーワード入力時に検索用データベースに格納されているネットワーク情報が情報検索装置（検索サーバー）によって検索され、情報検索システムからユーザーに必要なネットワーク情報が提供される。

【0004】また、他の情報検索システムでは、情報検索システムがネットワーク情報を収集するのではなく、必要なネットワーク情報を各ユーザーが検索用データベースに登録する。検索キーワードが入力された場合、上述したように、入力されたキーワードを基にして検索用データベースに格納されているネットワーク情報が情報検索装置によって検索され、情報検索システムからユーザーに必要な情報が提供される。

【0005】従って、ユーザーによる検索キーワードの入力毎に情報検索装置によって各コンピュータシステム内のネットワーク情報が検索されるわけではない。大規模

ネットワークには多数のコンピュータシステムが接続されているので、各コンピュータシステム内の全てのネットワーク情報を検索するのに長時間を要するからである。

【0006】ネットワーク情報の検索は種々の検索ツールを用いて行われているが、例えばYahoo!、Lycos 等のWWW (World Wide Web) におけるサーチエンジンというプログラムを用いてネットワーク情報の検索を行う場合、ユーザは次のような操作を行う。

【0007】Yahoo!では、各ユーザの登録によって検索用データベースに格納されているネットワーク情報は階層構造になっている。従って、Yahoo!のホームページ上に表示されているカテゴリ分けされた項目の中から所定の項目をユーザが選択する。これにより選択したカテゴリに関連するサブカテゴリが表示され、さらにユーザが所定のサブカテゴリを選択する。このような対話形式の選択を繰返して行うことによって何階層か下って最終的に個々のネットワーク情報の見出しが表示されることになる。

【0008】また、Yahoo!では、検索キーワード（検索文字列）を入力することによってネットワーク情報を自動検索する。基本的な検索では、検索キーワードを入力して検索ボタンを押せば、画面上に検索結果が表示される。検索結果に満足できなければ、オプションボタンを押すことにより検索方法を変更する。すなわち、オプション選択では、（1）検索範囲を題目、URL (Universal Resource Locator、WWWにおけるページを特定するための識別子)、テキストのどれかに限定する指定、

（2）キーワードの大文字、小文字を区別する指定、

（3）キーワードが全体一致かまたは部分一致かを選択する指定、（4）複数のキーワードを入力する場合に全てのキーワードを含んでいるものかまたは少なくとも1つのキーワードを含んでいればよいものか（ブール論理による表記）を選択する指定、（5）検索結果の数を制限するかまたは全てを検索するかの指定が可能であり、これらの指定を適時組合せて検索を行う。

【0009】Lycos では、スパイダー、ワーム、ロボット等の名前で呼ばれているプログラムを用いて自動的にネットワーク情報を収集し、これを検索用データベースに格納する。ネットワーク情報の検索は検索キーワードの入力により上述したYahoo!と同様な操作で行われる。

【0010】以上のように、ユーザ各個人に適した内容でかつ適した分量のネットワーク情報をユーザに提供するためには、各ユーザが求め必要なネットワーク情報に関するキーワード等を入力し、入力されたキーワードを基にしてユーザ各個人がどのようなネットワーク情報を必要としているかを情報検索システムが判断し、この判断を基にして必要なネットワーク情報をユーザに提供していた。

【0011】

【発明が解決しようとする課題】しかし、情報検索装置を用いてネットワーク情報を検索する場合、上述した対話形式では検索に時間がかかるという問題があった。また、検索キーワードの入力による自動検索では、各ユーザによって必要なネットワーク情報は異なるので、各ユーザ毎に適切な検索ツールの選択、適切な検索キーワードの入力等の初期設定を行わなければならないが、特に適切な検索キーワードの入力は容易ではないため、ユーザに大きな負担がかかるという問題があった。

10 【0012】本発明は上記事情によってなされたものであり、本発明の目的はキーワード等をユーザに能動的に設定されることなく、大規模ネットワークに接続しているコンピュータシステム（サーバー）内のネットワーク情報にユーザがアクセスした場合のアクセス記録情報を参照および解析し、大規模ネットワークに接続しているコンピュータシステム内のどのようなネットワーク情報をユーザが必要としているかを判断し、この判断を基にして必要なネットワーク情報をユーザに提供するための情報検索方法、情報検索システム、情報検索端末装置および情報検索装置を提供することである。

【0013】

【課題を解決するための手段】上記の課題を解決するため、本発明の情報検索システムは、ネットワークと、このネットワークに接続している複数のコンピュータシステムと、各コンピュータシステム内のネットワーク情報を格納するネットワーク情報格納手段と、アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納する手段と、アクセス記録情報を基にして頻繁にアクセスしたネットワーク情報の題目を格納する参照先情報格納手段と、ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納する手段と、参照先情報格納手段に格納されているネットワーク情報の題目から重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、決定した情報分野を基にしてネットワーク情報格納手段に格納されているネットワーク情報の中から所定のネットワーク情報を選択する手段と、選択したネットワーク情報を提供する手段とを有することを特徴とする。

30 40 【0014】また、本発明の情報検索方法は、ネットワークに接続している複数のコンピュータシステム内のネットワーク情報を格納し、アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納し、アクセス記録情報を基にして頻繁にアクセスしたネットワーク情報のアドレスとその題目を格納し、ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納し、格納されているネットワーク情報の題目から重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、決定した情報分野を基にして格

納しているネットワーク情報の中から所定のネットワーク情報を選択し、選択したネットワーク情報を提供することを特徴とする。

【0015】さらに、本発明の情報検索端末装置は、アドレスを指定してネットワークに接続している複数のコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報を格納する手段と、格納したアクセス記録情報を基にして頻繁にアクセスしたネットワーク情報の題目を格納する手段と、ネットワーク情報の情報分野およびそれに属する単語を対応づけて格納する手段と、格納したネットワーク情報の題目から重要度の高い単語を抽出し、抽出した重要度の高い単語が属するネットワーク情報の情報分野を決定し、決定した情報分野を基にして格納されているネットワーク情報の中から所定のネットワーク情報を選択し、選択したネットワーク情報を取得する手段を有することを特徴とする。

【0016】さらにまた、本発明の情報検索装置は、ネットワークに接続している複数のコンピュータシステム内のネットワーク情報を格納するネットワーク情報格納手段と、アドレスを指定してコンピュータシステム内のネットワーク情報をアクセスした場合のアクセス記録情報の中で頻繁にアクセスされたネットワーク情報の題目から重要度の高い単語が抽出され、予め単語に対応付けられた情報分野を参照して重要度の高い単語が属する情報分野が決定されることによって、この情報分野を基にしてネットワーク情報格納手段に格納しているネットワーク情報の中から選択された所定のネットワーク情報を提供する手段とを有することを特徴とする。

【0017】

【発明の実施の形態】以下、本発明の実施の形態を図面を参照して詳細に説明する。

【0018】図1は本発明の実施の一形態の情報検索システムの構成を示す図である。図1に示す情報検索システムは、ワークステーション(情報検索端末装置)10、大規模ネットワーク11、情報検索装置(検索サーバー)12、記憶装置13およびコンピュータシステム(サーバー)14~17によって構成されている。ワークステーション10は、表示装置1、通信装置2、入力装置3、制御装置4、外部記憶装置5および時計装置6によって構成されている。情報検索装置12には記憶装置13が接続されている。ワークステーション10は、インターネットのような大規模ネットワーク11を介して情報検索装置12やコンピュータシステム14~17に接続されている。

【0019】通信装置2は大規模ネットワーク11に接続されている情報検索装置12やコンピュータシステム14~17内の情報(ネットワーク情報)にアクセスする。このネットワーク情報には、テキスト情報、画像情報、音声情報等の少なくとも1つの情報が含まれる。

【0020】制御装置4は、通信装置2を介して大規模

ネットワーク11に接続されているコンピュータシステム14~17内のネットワーク情報を収集し、収集したネットワーク情報を外部記憶装置5に格納させ、大規模ネットワーク11を通してユーザがアクセスしたコンピュータシステム内のネットワーク情報のアドレス、アクセス日時、アクセス回数等の個人アクセス記録情報を外部記憶装置5に格納させる。また、制御装置4は、ユーザの個人参照先情報、ユーザの個人アクセス記録情報を解析し、その解析結果を基にして情報検索装置12に接続された記憶装置13からユーザに適したネットワーク情報を取得し、そのネットワーク情報を表示装置1に表示されること等によりユーザに提供する。

【0021】入力装置3は、キーボード、マウス等からなり、大規模ネットワーク11に接続されているコンピュータシステム内のネットワーク情報のアドレスや各種指示を入力する。

【0022】外部記憶装置5は、ハードディスク等によって構成され、個人単語重要度情報、情報分野情報、個人参照先情報、および個人アクセス記録情報を格納する。

【0023】表示装置1は、カラーCRTディスプレイ、カラー液晶ディスプレイ等からなり、入力装置3から入力されたネットワーク情報のアドレスや各種指示を表示し、また、種々の情報取得のためのメニューやネットワーク情報の検索結果等を表示する。

【0024】情報検索装置12は、各コンピュータシステム14~17内のネットワーク情報を定期的に収集し、収集したネットワーク情報を検索用データベースとして用いられる記憶装置13に格納する。ユーザによって情報検索のための検索キーワードが入力装置3から入力された場合、入力されたキーワードを基にしてキーワード入力時に記憶装置13に格納されているネットワーク情報が情報検索装置12によって検索され、その検索結果を表示装置1に表示することによりユーザに必要なネットワーク情報が提供される。

【0025】例えばYahoo!を用いてネットワーク情報の検索を行う場合、検索キーワードを入力する。基本的な検索では、検索キーワードを入力して検索ボタンを押せば、表示装置1に検索結果が表示される。表示された検索結果に満足できなければ、オプションボタンを押すことにより検索方法を変更する。すなわち、オプション選択では、検索範囲を題目、URL、テキストのどれかに限定する指定、キーワードの大文字、小文字を区別する指定、キーワードが全体一致かまたは部分一致かを選択する指定、複数のキーワードを入力する場合に全てのキーワードを含んでいるものかまたは少なくとも1つのキーワードを含んでいればよいものか(プール論理による表記)を選択する指定、検索結果の数を制限するかまたは全てを検索するかの指定が可能であり、これらの指定を適時組合せて検索を行う。

【0026】時計装置6は、水晶発振器等からなり、一定時間ごとに割込み処理を行い、現在の時刻を図示しないメモリ中の日時バッファに格納する。

【0027】図2は情報検索装置12および記憶装置13の構成を示す図であり、図3は制御装置4および外部記憶装置5の構成を示す図である。

【0028】情報検索装置12は情報データベース作成部12aおよび情報データベース検索部12bからなる。また、記憶装置13は大規模ネットワーク情報データベース13aを有する。

【0029】制御装置4は、個人情報選択部4aおよび個人参照先情報データベース作成部4bからなる。また、外部記憶装置5は、個人単語重要度情報データベース5a、情報分野情報データベース5b、個人参照先情報データベース5c、および個人アクセス記録情報データベース5dを有する。

【0030】情報データベース作成部12aは、大規模ネットワーク11に接続されているコンピュータシステム内のネットワーク情報を収集し、収集したネットワーク情報を大規模ネットワーク情報データベース13aに格納する。情報データベース検索部12bは、大規模ネットワーク情報データベース13aに格納されているネットワーク情報の中でユーザ個人に適したネットワーク情報を選択する。

【0031】個人参照先情報データベース作成部4bは、ユーザが大規模ネットワーク11に接続しているコンピュータシステム内のネットワーク情報をアクセスした記録を示すアクセス記録情報を個人アクセス記録情報データベース5dに格納させ、このアクセス記録情報を参照してユーザが頻繁にアクセスしているコンピュータシステム内のネットワーク情報の所在アドレスとその情報題目を個人参照先情報データベース5cに格納させる。

【0032】個人情報選択部4aは、個人参照先情報データベース5cを解析して個人単語重要度情報データベース5aを作成し、個人単語重要度情報データベース5aや情報分野情報データベース5bを参照し情報データベース検索部12bを通してユーザに適したネットワーク情報を選択し、選択したネットワーク情報をユーザに提供する。

【0033】次に、図1に示す情報検索システムの処理動作について図11のフローチャートを参照して説明する。

【0034】情報データベース作成部12aは、大規模ネットワーク11に接続されているコンピュータシステム内のネットワーク情報を収集し、収集したネットワーク情報の所在アドレス、情報題目(情報見出し)および情報本文(情報本体)からなる大規模ネットワーク情報データベース13aを作成する(ステップA1)。大規模ネットワーク情報データベース13aの構造の一例を

図4に示す。なお、このデータベース作成処理は毎日行われる。

【0035】個人参照先情報データベース作成部4bは、ユーザが大規模ネットワーク11に接続されているコンピュータシステムのアドレスを指定して入力装置3の図示しないネットワーク通信ボタンを選択した場合、このコンピュータシステムの指定したアドレスにアクセスしてそのアドレスに所在するネットワーク情報を取得し(ステップA2)、取得したネットワーク情報を表示装置1に表示させる。その際、大規模ネットワーク11に接続されているコンピュータシステム内のネットワーク情報をユーザが取得したこと示す個人アクセス記録情報が個人アクセス記録情報データベース5dに格納される(ステップA3)。

【0036】図7および図8に示すように、個人アクセス記録情報には、ネットワーク情報の情報ID、情報題目、情報の所在アドレス、最終アクセス日時、総アクセス回数、例えば最近1か月間にアクセスした回数を示す一定期間アクセス回数、テキスト情報を有するファイルのファイル名、画像情報を有するファイルのファイル名、音声情報を有するファイルのファイル名が含まれる。

【0037】また、個人参照先情報データベース作成部4bは、上記の個人アクセス記録情報を参照して個人参照先情報の記憶処理を行う。この個人参照先情報の記憶処理により、ユーザが頻繁にアクセスする情報題目とそのアドレスを対応付けて作成した個人参照先情報データベース5c(図5参照)がローカルディスク等の外部記憶装置5に格納される。個人参照先情報データベース5cに格納されている1つのネットワーク情報の情報題目が入力装置3から入力された選択指示により選択された場合、通信装置2を通して大規模ネットワーク11に接続されているコンピュータシステムにおけるそのネットワーク情報が存在するアドレスに直接通信が行われ、これにより、そのネットワーク情報を取得できる。

【0038】個人参照先情報データベース5cは次のような2通りの方法で求め作成される。第1の方法では、ユーザが頻繁にアクセスするアドレスを記憶する場合、実際にそのアドレスをアクセスしている時にユーザが入力装置3の図示しない記憶ボタンを選択し、そのアドレスと情報題目を個人参照先情報データベース5cに格納しそれを作成する。第2の方法では、上記個人アクセス記録情報(図7および図8参照)を定期的に参照して解析し、その解析結果を基にして個人参照先情報データベース5cを更新して作成する(ステップA3)。この第2の方法では、所定期間内に所定回数以上アクセスしたネットワーク情報のアドレスとその情報題目が自動的に個人参照先情報データベース5cに登録される一方、個人参照先情報データベース5cに既に登録されているネットワーク情報の中でアクセス回数が所定期間内に所定

回数以下であるネットワーク情報のアドレスとその情報題目が個人参照先情報データベース5cから削除される。

【0039】なお、個人参照先情報データベース5cはユーザによる入力装置3からの編集指示により削除等の編集が可能である。また、個人参照先情報データベース5cに格納されているネットワーク情報の中で関連性の高い複数のネットワーク情報をまとめてユーザ固有の項目名（中項目）を付けて格納し、さらに複数の中項目をまとめて大項目を付けて格納する等、個人参照先情報データベース5cを分類された複数の階層からなる階層構造にことができる。このような階層構造を有する個人参照先情報データベース5cの構造の一例を図6に示す。

【0040】以上の処理後、個人情報選択部4aは、ユーザ各個人に適したネットワーク情報を提供するための設定を図12に示すフローチャートに従って行う（ステップA4）。この設定に際しては、予め作成された情報分野の項目とその情報分野に関連するキーワードの項目を対応付けして格納している情報分野情報データベース5bが用いられる。情報分野情報データベース5bの構造の一例を図10に示す。

【0041】まず、後述するユーザの情報評価を基にして個人単語重要度情報データベース5aが更新されたかどうかが判定される（ステップB1）。個人単語重要度情報データベース5aが更新されている場合、更新された個人単語重要度情報データベース5aを基にしてネットワーク情報が選択され、選択されたネットワーク情報がユーザに提供される（ステップB2）。

【0042】個人単語重要度情報データベース5aが更新されていない場合、個人情報選択を簡易版、標準版、または詳細版のいずれにするかをユーザに選択させる（ステップB3）。

【0043】個人情報選択を簡易版にする場合（ステップB4）、図13に示すようなフローチャートに従って簡易個人情報選択が行われる（ステップB5）。すなわち、図6に示されているような個人参照先情報データベース5cに格納されている階層構造の項目（大項目、中項目）や情報題目（情報見出し）を参照し、次のような解析を行う（ステップC1）。

【0044】すなわち、項目や情報題目に含まれる単語が英語の場合にはスペースで区切ることにより項目や情報題目から単語を抽出し、日本語の場合には形態素解析により単語を抽出する。単語抽出の際には、抽出された単語にその所在場所により重要度点を付ける。例えば、個人参照先情報データベース5cの大項目に含まれる単語は10点、中項目に含まれる単語は3点、情報項目に含まれる単語は3点とする。そして、各単語が出現する度に重要度点を加算して各単語における最終的な重要度点を取得し、図9に示すような単語と対応する重要度点

を個人単語重要度情報データベース5aに格納してそれを作成する（ステップC2）。

【0045】個人情報選択を標準版にする場合（ステップB6）、図14に示すようなフローチャートに従って標準個人情報選択が行われる（ステップB7）。すなわち、ステップC1およびC2の処理と同様に、個人参照先情報データベース5cに格納されている階層構造の項目（大項目、中項目）や情報題目（情報見出し）を参照して解析し（ステップD1）、単語を抽出して各単語に重要度点を付けた後、単語と対応する重要度点を個人単語重要度情報データベース5aに格納してそれを作成する（ステップD2）。

【0046】さらに、個人参照先情報データベース5cに登録されているアドレスに所在するネットワーク11に接続されているコンピュータシステム内のネットワーク情報本体のテキストをアクセスして取得し、このテキストに対して次のような解析を行う（ステップD3）。

【0047】取得したテキストに含まれる単語が英語の場合にはスペースで区切ることによりそのテキストから単語を抽出し、日本語の場合には形態素解析により単語を抽出する。単語抽出の際には、抽出された単語にその所在場所により重要度点を付ける。例えば、そのテキストの見出しに含まれる単語は10点、中見出しに含まれる単語は5点、小見出しに含まれる単語は3点、本文中の単語は1点とする。そして、各単語が出現する度に重要度点を加算して各単語における最終的な重要度点を取得し、取得した重要度点を個人単語重要度情報データベース5aに既に格納されている各単語の重要度点に加算して得られた重要度点を個人単語重要度情報データベース5aに格納してそれを更新する（ステップD4）。

【0048】個人情報選択を詳細版にする場合（ステップB8）、図15に示すようなフローチャートに従って詳細個人情報選択が行われる（ステップB9）。すなわち、ステップD1～D4の処理と同様に、個人参照先情報データベース5cに格納されている階層構造の項目や情報題目を参照して解析し（ステップE1）、単語を抽出して各単語に重要度点を付けた後、単語と対応する重要度点を個人単語重要度情報データベース5aに格納してそれを作成する（ステップE2）。

【0049】さらに、個人参照先情報データベース5cに登録されているアドレスに所在するネットワーク11に接続されているコンピュータシステム内のネットワーク情報本体のテキストをアクセスして取得し、このテキストに対して次のような解析を行う（ステップE3）。

【0050】取得したテキストに含まれる単語が英語の場合にはスペースで区切ることにより取得したテキストから単語を抽出し、日本語の場合には形態素解析により単語を抽出する。単語抽出の際には、抽出された単語にその所在場所により重要度点を付ける。例えば、そのテキストの見出しに含まれる単語は10点、中見出しに含

まれる単語は5点、小見出しに含まれる単語は3点、本文中の単語は1点とする。そして、各単語が出現する度に重要度点を加算して各単語における最終的な重要度点を取得し、取得した重要度点を個人単語重要度情報データベース5aに既に格納されている各単語の重要度点に加算して得られた重要度点を個人単語重要度情報データベース5aに格納してそれを更新する(ステップE4)。

【0051】また、個人参照先情報データベース5cに登録されているアドレスに所在するネットワーク情報本体から参照している他のネットワーク情報のテキストを例えば3つ目の深さまで参照先を辿って取得し、次のような解析を行う(ステップE5)。

【0052】取得したテキストに含まれる単語が英語の場合にはスペースで区切ることによりそのテキストから単語を抽出し、日本語の場合には形態素解析により単語を抽出する。単語抽出の際には、各深さの参照先のネットワーク情報のテキストから抽出される単語にその所在場所により重要度点を付ける。例えば、1つ目の深さの参照先のネットワーク情報のテキストの見出しに含まれる単語は10点、中見出しに含まれる単語は5点、小見出しに含まれる単語は3点、本文中の単語は1点とし、また、2つ目の深さの参照先のネットワーク情報のテキストの見出しに含まれる単語は5点、中見出しに含まれる単語は2.5点、小見出しに含まれる単語は1.5点、本文中の単語は0.5点とし、さらに、3つ目の深さの参照先のネットワーク情報のテキストの見出しに含まれる単語は3点、中見出しに含まれる単語は1.5点、小見出しに含まれる単語は1点、本文中の単語は0.3点とする。そして、各単語が出現する度に重要度点を加算して各単語における最終的な重要度点を取得し、取得した重要度点を個人単語重要度情報データベース5aに既に格納されている各単語の重要度点に加算して得られた重要度点を個人単語重要度情報データベース5aに格納してそれを更新する(ステップE6)。

【0053】簡易版、標準版、または詳細版のいずれかの個人情報選択に従って個人単語重要度情報データベース5aを作成した後、この個人単語重要度情報データベース5aを参照して重要度点が高い単語から順にソートを行う。このソート後、重要度点の高い例えば上位3つの単語(図9では、AAA、BBB、およびCCC)を用いて重要度点の高い単語の順序で情報分野情報データベース5b(図10参照)に格納されているキーワード項目を検索する(ステップC3、ステップD5、ステップE7)。

【0054】単語にマッチしたキーワードが情報分野情報データベース5bに格納されている場合、そのキーワードが属する情報分野の項目をユーザに適した情報分野であると決定する。例えば、単語にマッチしたキーワードが体操であれば、情報分野はスポーツに決定される。

なお、複数の情報分野でマッチした場合、重要度点のより高い単語にマッチしたキーワードが属する情報分野を優先し、その情報分野をユーザに適した情報分野であると決定する。

【0055】情報分野が決定した後、情報データベース検索部12bでは、決定した情報分野に属するキーワード項目中の複数のキーワードをORで連結した検索式を用いて大規模ネットワーク情報データベース13aを検索する(ステップC4、ステップD6、ステップE8)。

10 例えれば、決定した情報分野がスポーツであれば、体操、水泳、サッカー、オリンピック等のキーワードを用いてOR検索を行う。この検索によって得られたネットワーク情報は、定期的に例えば1日に1回電子メール等によってユーザに提供される(ステップC5、ステップD7、ステップE9)。

【0056】ユーザに提供されるネットワーク情報には2つの形式がある。1つはテキストのみのファイル、もう1つはテキスト、画像情報、および音声情報を含むSGML(Standard Generalized Markup Language)形式のファイルである。後者の形式のファイルを選択する場合、SGML文書表示ソフトウェア(例えばNetscape社が販売しているNetscape Navigator等)を用いてそのネットワーク情報が表示装置1に表示される。

【0057】表示装置1に表示されたネットワーク情報は図16に示すフローチャートに従ってユーザによって評価される(ステップA5)。まず、ユーザに提供するネットワーク情報があるかどうかが判断され(ステップF1)、ユーザに提供するネットワーク情報がある場合には表示装置1にそのネットワーク情報が表示される(ステップF2)。

【0058】提供されたネットワーク情報がユーザに必要なネットワーク情報である場合にはユーザが入力装置3の図示しない肯定ボタンを押し、そのネットワーク情報が不要である場合にはユーザが入力装置3の図示しない否定ボタンを押すことにより、そのネットワーク情報が評価される(ステップF3)。

【0059】ユーザが肯定ボタンを押すことによりそのネットワーク情報が肯定評価された場合、その時に表示装置3に表示されているネットワーク情報のテキストを40 解析してそのテキストから単語を順次抽出し(ステップF4)、各単語にその所在位置により前述した重み付けて重要度点を付け、個人単語重要度情報データベース5aを更新する(ステップF5)。例えば、テキストの見出し中の単語は10点、強調文字で示された単語は5点、本文は1点とし、各単語が出現する度に重要度点を加算して個人単語重要度情報データベース5aを更新する。このような情報評価は提供されているネットワーク情報の全てについて繰返し行われる。

【0060】その後、定期的に更新される個人単語重要度情報データベース5aを参照して重要度点が高い単語

から順にソートを行う。このソート後、重要度点の高い例えば上位3つの単語を用いて重要度点の高い単語の順序で情報分野情報データベース5bに格納されているキーワード項目を検索する。

【0061】単語にマッチしたキーワードが情報分野情報データベース5bに格納されている場合、そのキーワードが属する情報分野の項目をユーザに適した情報分野であると決定する。複数の情報分野でマッチした場合、重要度点の高い単語にマッチしたキーワードが属する情報分野を優先し、その情報分野をユーザに適した情報分野であると決定する。

【0062】情報分野が決定した後、情報データベース検索部12bでは、決定した情報分野に属するキーワード項目中の複数のキーワードをORで連結した検索式を用いて大規模ネットワーク情報データベース13aを検索する。この検索により得られたネットワーク情報は定期的にユーザに提供される。

【0063】さらに、例えば1日に1回個人参照先情報データベース5cに格納されているアドレスに所在するネットワーク情報が更新されているかどうかをチェックし、それが更新されている場合にはその旨を例えば表示装置1に表示することによってそれをユーザに知らせる。

【0064】本発明は上記実施の形態に限定されることなく、本発明の要旨の範囲内において種々の変形実施が可能である。本発明の実施の形態では、1つのワークステーション内で処理が行われるが、処理の負担を軽減させるためには、サーバクライアント型で処理を行うことが可能である。本発明は、文書検索装置、情報検索装置、ネットワーク情報検索装置等に応用可能である。

【0065】

【発明の効果】以上、詳述したように、本発明によれば、ユーザに必要なネットワーク情報をユーザに提供するための設定は、ユーザにキーワードを入力させることなく、ユーザが大規模ネットワークに接続されているコンピュータシステム（サーバー）に対してアクセスした時に作成される個人参照先情報データベースを解析することにより行うことができる。すなわち、ユーザによるネットワーク情報のアクセスの初期の段階では、ユーザがキーワード等を入力することによってユーザ自身が必要なネットワーク情報を大規模ネットワークに接続されているコンピュータシステムに対してアクセスする必要があるが、ネットワーク情報のアクセスが頻繁に行われる段階においてユーザがアクセスするネットワーク情報が特定の情報分野に固定されている場合、ユーザがキーワード等を入力しなくとも、その情報分野に関するネットワーク情報をユーザに提供することができる。従つて、ネットワーク情報のアクセスを行うごとにユーザがキーワードを入力するという負担を軽減することができる。

【0066】また、個人重要単語情報データベースを用いてユーザに必要なネットワーク情報を提供するための初期設定として、個人参照先情報データベースに格納されているネットワーク情報のみで行う簡易設定、簡易設定で必要なネットワーク情報のみならずコンピュータシステム内のネットワーク情報を用いて行う標準設定および詳細設定の3つの設定方法を準備しているので、緊急度に応じて簡易版、標準版、詳細版のいずれかで初期設定することができる。

10 【0067】さらに、ユーザが能動的に頻繁にアクセスするアドレスを個人参照先情報データベースに格納するだけでなく、ユーザが所定期間中に実際にアクセスする頻度が高いアドレスを自動的に個人参照先情報データベースに格納し、さらに以前はアクセス頻度が高くても時間の経過とともにアクセス頻度が低くなったアドレスを個人参照先情報データベースから自動的に削除することにより、アクセス頻度が高いアドレスのみを個人参照先情報データベースに格納できる。

【図面の簡単な説明】

20 【図1】本発明の情報検索システムの概略構成を示す図。

【図2】図1に示すシステムの情報検索装置および記憶装置の構成を示す図。

【図3】図1に示すシステムの制御装置および外部記憶装置の構成を示す図。

【図4】大規模ネットワーク情報データベースの構造の例を示す図。

【図5】個人参照先情報データベースの構造の例を示す図。

30 【図6】個人参照先情報データベースの構造の他の例を示す図。

【図7】個人アクセス記録情報データベースの構造の例を示す図。

【図8】個人アクセス記録情報データベースの構造の例を示す図。

【図9】個人単語重要度情報データベースの構造の例を示す図。

【図10】情報分野情報データベースの構造の例を示す図。

40 【図11】図1に示す情報検索システムの処理動作を示すフローチャート。

【図12】個人情報設定を示すフローチャート。

【図13】簡易個人情報選択を示すフローチャート。

【図14】標準個人情報選択を示すフローチャート。

【図15】詳細個人情報選択を示すフローチャート。

【図16】ユーザ情報評価および個人単語重要度情報データベースの更新を示すフローチャート。

【符号の説明】

1 表示装置

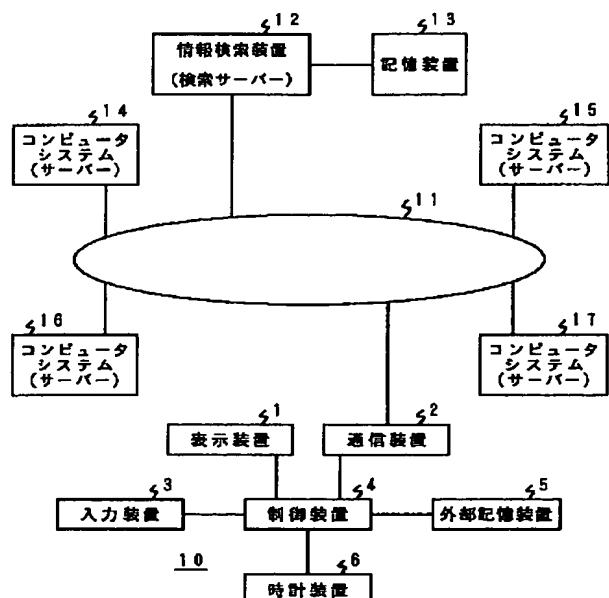
2 通信装置

3 入力装置
 4 制御装置
 4 a 個人情報選択部
 4 b 個人参照先情報データベース作成部
 5 a 個人単語重要度情報データベース
 5 b 情報分野情報データベース
 5 c 個人参照先情報データベース
 5 d 個人アクセス記録情報データベース
 5 外部記憶装置
 6 時計装置

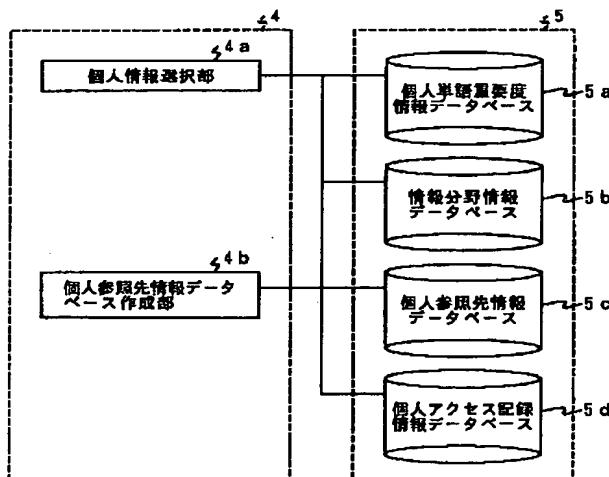
* 10 ワークステーション
 11 大規模ネットワーク
 12 情報検索装置
 12 a 情報データベース作成部
 12 b 情報データベース検索部
 13 記憶装置
 13 a 大規模ネットワーク情報データベース
 14、15、16、17 コンピュータシステム（サーバー）

* 10

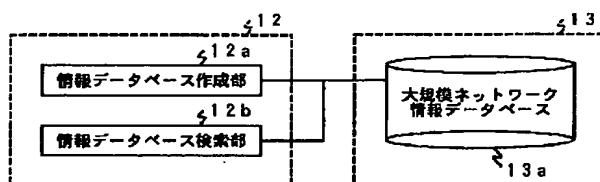
【図1】



【図3】



【図2】



【図4】

アドレス	情報見出し	情報本体
AAAA	A	Aというのは、情報に関する_____
BBBB	B	Bというのは、_____
-----	-----	-----
-----	-----	-----

【図5】

情報題目	アドレス
A新聞のホームページ	http://www.aaa.co.jp/...
B新聞のホームページ	http://www.bbb.co.jp/...
C新聞のホームページ	http://www.ccc.co.jp/...
D社のホームページ	http://www.ddd.co.jp/...
E社のホームページ	http://www.eee.co.jp/...
F社のホームページ	http://www.fff.co.jp/...
「G H I J」について	http://www.fff.co.jp/...
K K さんについて	http://www.kkk.co.jp/...
L社のホームページ	http://www.lll.co.jp/...
M社のホームページ	http://www.mmm.co.jp/...

【図10】

分野	キーワード
スポーツ	体操、水泳、サッカー、オリンピック、.....
政治	永田町、自民党、社会党、国会、参議院、選挙、.....

【図6】

大項目	中項目	情報題目	アドレス
ニュース	一般	A新聞のホームページ	http://www.aaa.co.jp/...
		B新聞のホームページ	http://www.bbb.co.jp/...
ソフト	WWW	C新聞のホームページ	http://www.ccc.co.jp/...
		D社のホームページ	http://wwwddd.co.jp/...
	その他	E社のホームページ	http://www.eee.co.jp/...
個人的	映画	F社のホームページ	http://www.fff.co.jp/...
		「GHIJ」について	http://www.ggg.co.jp/...
	音楽	KKさんについて	http://www.ddd.co.jp/...
		L社のホームページ	http://www.fff.co.jp/...
		M社のホームページ	http://www.fff.co.jp/...

【図7】

情報ID	情報題目	アドレス
1	A新聞のホームページ	http://www.aaa.co.jp/...
2	B新聞のホームページ	http://www.bbb.co.jp/...
3	C新聞のホームページ	http://www.ccc.co.jp/...
4	D社のホームページ	http://wwwddd.co.jp/...
5	E社のホームページ	http://www.eee.co.jp/...
6	F社のホームページ	http://www.fff.co.jp/...
7	「GHIJ」について	http://www.ggg.co.jp/...
8	KKさんについて	http://www.ddd.co.jp/...
9	L社のホームページ	http://www.fff.co.jp/...
---	---	---

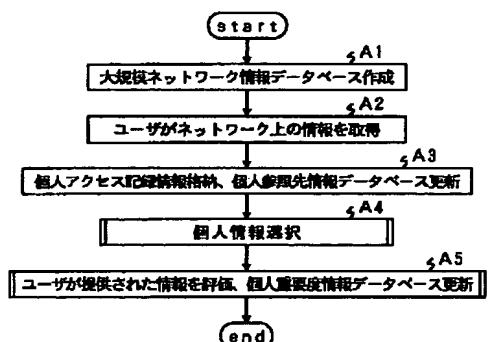
【図8】

(上記表の続き)						
情報ID	最終アクセス日時	最終アクセス回数	一定期間アクセス回数	テキスト	画像	音声
1	96/5/23 14:23	123	1	aaa.txt	aaa.gif	aaa.wav
2	96/5/3 8:24	33	10	bbb.txt	bbb.gif	bbb.wav
3	96/5/11 14:23	44	9	ccc.txt	ccc.gif	ccc.wav
4						
5						
6						
7						
8						
9						

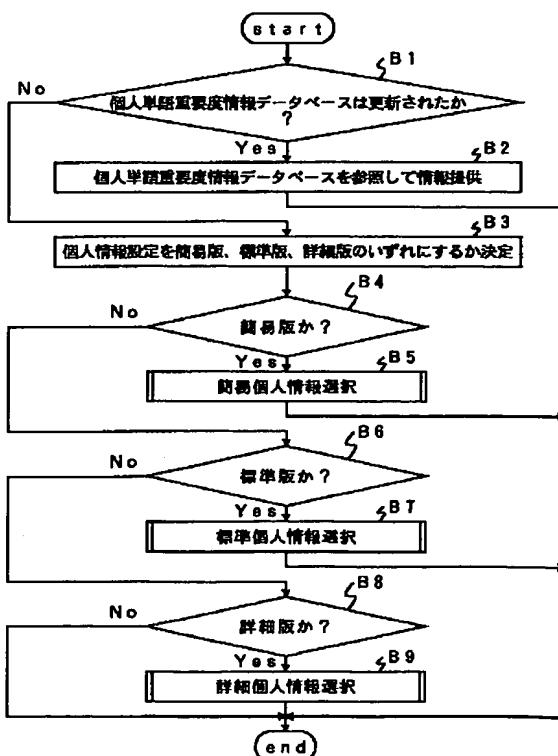
【図9】

単語	重要度点
AAA	78
BBB	56
CCC	45
DDD	33
EEE	31
FFF	20
GGG	17
HHH	16
III	14
---	---

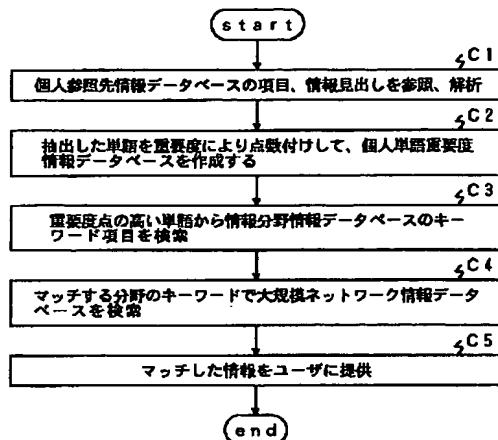
【図11】



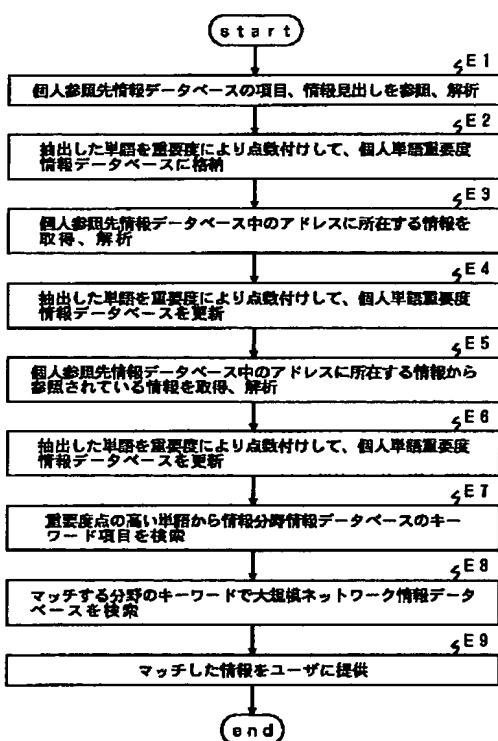
【図12】



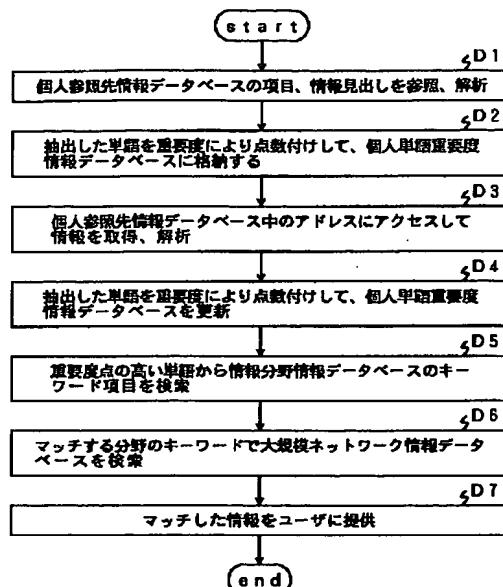
【図13】



【図15】



【図14】



【図16】

